

Pour une évaluation subversive

THOMAS LEMAIGRE ET CHRISTOPHE MINCKE

C'est devenu une antienne, il faut évaluer. Évaluer quoi ? Mais évaluer, pardi ! Tout ! Tous ! Tout le temps !

Notre société aime la qualité, de même que les performances. Elle promet donc l'évaluation continue. Le *monitoring*, pour être plus précis (mais moins francophone), c'est-à-dire un suivi continu de l'activité en fonction de variables mesurables. Autrefois, le contrôle intervenait essentiellement par le biais de coups de sondes : enquêtes, procédures d'évaluation, examens, etc. Mais l'informatique, entre autres, nous permet désormais de connaître en permanence la position des trains et camions, le flux de voitures en un point, la productivité de chaque travailleur dans chacune des tâches qui lui sont assignées, les délais de réponse, les temps d'attente, les taux de satisfaction, le suivi des programmes de formation, les indices d'écoute, etc.

La tentation est alors immense d'instaurer un *omnimonitoring* destiné à suivre chaque activité humaine en permanence, afin de traquer tout dysfonctionnement, tout goulot d'étranglement, toute rupture de charge, toute insatisfaction, toute baisse de performance. Du point de vue de la mise en œuvre, une difficulté se fait rapidement jour : on ne peut évaluer selon n'importe quel critère. Il n'existe pas de capteur à même de mesurer la joie des usagers d'un service ou l'épanouissement au travail, pas davantage l'humanité de la solution apportée à un problème. Il faut donc définir des critères objectivement mesurables au moyen des appareils rudimentaires que demeurent nos systèmes informatiques.

Dans certains cas, c'est particulièrement facile. Ainsi de l'évaluation des performances de nos transports en commun qui s'accommode sans difficulté de la comptabilisation des retards. Cet indicateur apparaît en effet à la fois objectif, aisément mesurable et consensuel... même si l'amabilité du contrôleur, la propreté des rames de

méto ou la qualité des infrastructures d'accueil ne se laissent pas si simplement capter. Potentiellement plus problématique est l'extension de ce type de logique au monde judiciaire et parajudiciaire, par exemple. La mesure de la charge de travail y est en effet devenue centrale dans l'évaluation du travail accompli. Mais qu'est-ce que rendre justice ? Gérer un flux de dossiers ou garantir le traitement de qualité d'une affaire ? Il en va de même dans les maisons de justice — qui gèrent, entre autres, le suivi des libérations conditionnelles — où un monitoring permanent des moindres actes des assistants de justice a été mis en place. D'où une question simple : peut-on considérer qu'accélérer les flux et réduire les goulots d'étranglement peut tenir lieu d'accroissement de la qualité ? La question a toute sa pertinence alors qu'il semble que soit relancée la réflexion sur une procédure de comparution immédiate en matière pénale.

On le devine déjà, un des dangers est celui de la contamination de secteurs toujours plus nombreux par une pure logique de gestion des flux. C'est ainsi que, régulièrement, on peut constater que la politique d'asile belge est présentée comme un succès du fait de l'accélération des flux de dossiers, de la réalisation d'économies et de l'augmentation du nombre de refus... On est en droit de se demander ce que ces critères peuvent bien nous dire de la qualité de cette politique. Sommes-nous bien certains qu'il faille se glorifier du refoulement d'une proportion croissante de candidats et de la baisse de la demande due à une réputation de pays peu accueillant ?

Dans le même ordre d'idées, on peut également s'interroger sur l'évaluation de notre enseignement. À cet égard, outre les traditionnelles plaintes du « monde de l'entreprise » sur l'inadéquation des formations, les enquêtes PISA font cycliquement parler d'elles. Un classement européen régulier et évaluant les performances par rapport à des indicateurs préétablis semble devenu le fin du fin en matière d'enseignement. On pourrait bien entendu s'interroger sur les raisons de la limitation de son évaluation à ces seuls critères, de même que sur la limitation du champ d'investigation aux enfants inscrits dans l'enseignement ordinaire, à l'exclusion de ceux qui en sont rejetés. Un enseignement inclusif sera à cet égard moins bien noté qu'un système qui ne se préoccupe que des plus doués et des mieux nantis.

Mais c'est un deuxième péril qui pointe ici : celui d'une redéfinition de l'activité évaluée dans le but d'améliorer les scores. C'est ainsi que la tâche des enseignants risque d'être redéfinie comme la préparation des élèves aux tests PISA. Ce serait bien naturel : si les professeurs sont blâmés pour les mauvais scores de leurs élèves, le plus simple pour eux serait de les entraîner à l'épreuve — plutôt que de les éduquer — et de rabattre ainsi sur quelques résultats limités les finalités multiples habituellement assignées au système éducatif. La partie pour le tout, le simple pour le complexe.

En réalité, nous voyons déjà se développer de nombreux exemples de ce type de démarche. Ce fut récemment le cas de la SNCB qui, pour améliorer ses évaluations, décida d'allonger ses temps de parcours, au point de faire descendre certains sous leur niveau de 1935... Les trains sont plus lents, mais l'évaluation s'améliore. Le monitoring est donc potentiellement porteur d'une dégradation de la qualité du service. La STIB, auparavant, avait déjà développé ce type de pratique, scindant des lignes de tram pour garantir la ponctualité sur la deuxième portion du trajet, même en cas de retard sur la première. Pour les voyageurs, la conséquence en est un allongement des temps de parcours et une évidente diminution du confort¹.

De là à penser que, par exemple, nos procédures judiciaires ou de décision en matière d'asile puissent être pensées au seul regard de la satisfaction de critères managériaux de gestion du contentieux, il n'y a qu'un pas... que le management public a franchi de longue date. La meilleure manière d'éviter un débat politique délicat est encore d'en limiter le champ par le recours à des outils de gestion, au premier rang desquels les systèmes d'information. Ceux-ci aideront en effet à reporter l'attention de questions complexes à quelques critères d'évaluation binaires.

La remise en question de ces systèmes d'information sera d'autant plus malaisée que personne, ou presque, n'en maîtrise le fonctionnement. À la limite, les décideurs deviennent des auxiliaires de dispositifs technologiques et managériaux sur la conception, le paramétrage et l'évolution desquels ils n'ont aucune prise directe — sans parler des travailleurs et des citoyens/usagers...

La puissance de calcul des machines, l'opacité des algorithmes, l'ingénierie infinie de la gestion des catégories, les oligopoles mondiaux des firmes qui maîtrisent et commercialisent les solutions informatiques limitent la maîtrise effective de ces outils à de rares spécialistes. Encore leur donne-t-on rarement les moyens d'investiguer en toute indépendance les outils mis en place et de diffuser largement leurs résultats. Le personnel dirigeant lui-même ne comprend que très peu les données dont il dispose et voit son rôle politique vidé d'une grande part de sa substance. Faut-il dès lors s'étonner de le voir réduit — de plus ou moins bon gré — à un rôle gestionnaire ?

Pourtant, des outils d'évaluation seraient extrêmement précieux pour soutenir la réflexion que nous devons impérativement mener sur l'avenir de nos sociétés. Mais, plus le temps passe, et plus il est difficile d'avouer que, depuis fort longtemps, plus personne n'y comprend rien et que le système n'est plus aux mains de personne². Et si l'on se rendait brusquement compte, au détour d'une réelle évaluation, que nous croyons agir depuis des années tout en ne faisant que nous débattre en vain ?

¹ Cocasse : le nouveau gouvernement bruxellois entend « évaluer » cette innovation.

² Voir notre dossier « Les tentations de l'impuissance », avril-mai 2014.

Dans ce contexte, nos gouvernements communautaires et régionaux, comme dans chaque déclaration de politique générale depuis quelques législatures, ont annoncé cet été leur intention d'évaluer nombre de mesures, de secteurs, de programmes. L'offre de formation qualifiante, les primes énergie, les aides de l'Awex, ou les clauses sociales dans les marchés publics en Wallonie; l'alliance emploi-environnement, la garantie jeunes, les chèques langues, les primes à la rénovation, ou encore le prêt logement en Région bruxelloise; et le décret « inscriptions », le décret « paysage de l'enseignement supérieur », la modularisation de l'enseignement qualifiant (CPU) ou encore le décret « éducation aux médias » en Fédération Wallonie-Bruxelles. À la Cocof, l'ambition évaluative est affichée de façon particulièrement volontariste: « De manière transversale, dans un souci constant d'amélioration des services aux citoyens, le Collège (NDLR: l'exécutif francophone bruxellois) s'engagera dans un processus d'évaluation permanent des politiques publiques. »

Tentation de l'omnimonitoring? Ambition velléitaire de tout réformer à la fois? Détricotage larvé des acquis de l'équipe précédente? Cache-sexe pour le manque d'idées politiques innovantes ou pour l'impossibilité de s'être mis d'accord dans le temps de négociations imparti? Fabrique à thermomètres que l'on mettra au rebut la fièvre venue? Arsenal à arguments à fourbir dans une succession de bras de fer désespérés avec le fédéral — pardon, avec le MR et la Flandre?

Ou encore, qui sait, réelle ambition de produire des connaissances partagées pour la décision et de nourrir le débat démocratique? Il faudrait, pour cela, que nous acceptions, avant que de procéder à des évaluations, de poser la question des valeurs qui les fondent... faute de quoi, il se pourrait bien que l'évaluation telle qu'elle se pratique aujourd'hui soit pire que son absence. ■